



# Bipolar Possibility Theory as a Basis for a Logic of Desire and Beliefs

Didier Dubois, Emiliano Lorini, Henri Prade

## ► To cite this version:

Didier Dubois, Emiliano Lorini, Henri Prade. Bipolar Possibility Theory as a Basis for a Logic of Desire and Beliefs. 7th International Conference on Scalable Uncertainty Management (SUM 2013), Sep 2013, Washington, DC, United States. pp.204-218, 10.1007/978-3-642-40381-1\_16 . hal-01212923

**HAL Id: hal-01212923**

**<https://hal.science/hal-01212923>**

Submitted on 7 Oct 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



## Open Archive TOULOUSE Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : <http://oatao.univ-toulouse.fr/>  
Eprints ID : 12779

Official URL: [http://dx.doi.org/10.1007/978-3-642-40381-1\\_16](http://dx.doi.org/10.1007/978-3-642-40381-1_16)

**To cite this version** : Dubois, Didier and Lorini, Emiliano and Prade, Henri  
*Bipolar Possibility Theory as a Basis for a Logic of Desire and Beliefs.*  
(2013) In: 7th International Conference on Scalable Uncertainty Management  
(SUM 2013), 16 September 2013 - 18 September 2013 (Washington, DC,  
United States).

Any correspondance concerning this service should be sent to the repository  
administrator: [staff-oatao@listes-diff.inp-toulouse.fr](mailto:staff-oatao@listes-diff.inp-toulouse.fr)

# Bipolar Possibility Theory as a Basis for a Logic of Desires and Beliefs

Didier Dubois, Emiliano Lorini, and Henri Prade

IRIT-CNRS, Université Paul Sabatier, Toulouse, France

**Abstract.** Bipolar possibility theory relies on the use of four set functions. On the one hand, a weak possibility and a strong necessity measure are increasing set functions, which are respectively max-decomposable with respect to union and min-decomposable with respect to intersection. On the other hand, strong possibility and weak necessity measures are two decreasing set functions, which are respectively min-decomposable with respect to union and max-decomposable with respect to intersection. In the first part of the paper we advocate the use of the last two functions for modeling a notion of graded desire. Moreover, we show that the combination of weak possibility and strong possibility allows us to model a notion of realistic desire, i.e., a desire that does not only account for satisfactoriness but also for its epistemic possibility. In the second part of the paper we show that possibility theory offers a semantic basis for developing a modal logic of beliefs and desires.

## 1 Introduction

Possibility theory has been originally proposed as an alternative approach to probability for modeling epistemic uncertainty, independently by two authors. In economics, Shackle [26] advocated a new view of the idea of expectation in terms of degree of surprise (a substitute for a degree of impossibility). Later in computer sciences, Zadeh [28] introduced a setting for modeling the information originated from linguistic statements in terms of fuzzy sets (understood as possibility distributions). Zadeh's proposal for a possibility theory relies on the idea of possibility measure, a max-decomposable set function w.r.t. union with values in  $[0, 1]$ . However, in these works, the duality between possibility and necessity (captured by a min-decomposable set function with respect to intersection) was not exploited. Later, it has been recognized that two other set functions, which contrast with the two previous ones by their decreasingness, also make sense in this setting [10]. These two latter set functions, which are dual of each other, model an idea of strong (guaranteed) possibility and of weak necessity respectively, while the original possibility measure that evaluates the consistency between the considered event and the available information, corresponds to a weak potential possibility.

The framework of possibility theory with its four basic set functions exhibits a rich structure of oppositions, which can be also closely related to other structures of oppositions that exist in modal logics and other settings such that formal concept analysis for instance [11]. Moreover, possibility theory is graded since the four set functions can take values in the unit interval. This very general setting can not only be interpreted in terms of uncertainty. It makes sense for preference modeling as well [2]. But it is also of interest when modeling situations that require modal logic languages, and where

grading modalities is meaningful. For instance, when modeling uncertainty, necessity measures are useful for representing beliefs and their epistemic entrenchments [9].

We provide here an investigation of the potentials of possibility theory for modeling the concept of desire. Indeed, although this concept has been already investigated in the past in artificial intelligence [19,20],<sup>1</sup> up to now, no clear connection between a theory of desires and possibility theory has been built. The rest of the paper is organized as follows. Section 2 presents a background on possibility theory. Section 3 discusses the modeling of desires in terms of strong possibility, as well the dual notion of potential desire in terms of weak necessity. We conclude by defining a notion of realistic desire, in the sense of desiring something that one considers epistemically possible. In Section 4 we introduce a modal logic of beliefs and desires based on possibility theory, more precisely, of realistic desires. The extension of this logic to graded desires is outlined. Finally, Section 5 points out some lines for further research on the relationship between possibility theory and the logic of emotions. A first version of sections 1-3 is in [22].

## 2 Background on Possibility Theory

Let  $\pi$  be a mapping from a set of worlds  $W$  to  $[0, 1]$  that rank-orders them. Note that this encompasses the particular case where  $\pi$  reduces to the characteristic function of a subset  $E \subseteq W$ . The possibility distribution  $\pi$  may represent a plausibility ordering (and  $E$  the available evidence) when modeling epistemic uncertainty, or a preference ordering ( $E$  is then the subset of satisfactory worlds) when modeling preferences. Let us recall the complete system of the 4 set functions underlying possibility theory [10] and their characteristic properties:

- i) The *(weak) possibility measure* (or potential possibility)  $\Pi(A) = \max_{w \in A} \pi(w)$  evaluates to what extent there is a world in  $A$  that is possible. When  $\pi$  reduces to  $E$ ,  $\Pi(A) = 1$  if  $A \cap E \neq \emptyset$ , which expresses the consistency of the event  $A$  with  $E$ , and  $\Pi(A) = 0$  otherwise. Possibility measures are characterized by the following decomposability property:  $\Pi(A \cup B) = \max(\Pi(A), \Pi(B))$ .
- ii) The dual *(strong or or actual) necessity measure*  $N(A) = \min_{w \notin A} 1 - \pi(w) = 1 - \Pi(\bar{A})$  evaluates to what extent it is certain (necessarily true) that all possible worlds are in  $A$ . When  $\pi$  reduces to  $E$ ,  $N(A) = 1$  if  $E \subseteq A$ , which expresses that  $E$  entails event  $A$  (when  $E$  represents evidence), and  $N(A) = 0$  otherwise. The duality of  $N$  w. r. t.  $\Pi$  expresses that  $A$  is all the more certain as the opposite event  $\bar{A}$  is impossible. Necessity measures are characterized by the following decomposability property:  $N(A \cap B) = \min(N(A), N(B))$ .
- iii) The *strong* (or actual, or “*guaranteed*”) possibility measure  $\Delta(A) = \min_{w \in A} \pi(w)$  evaluates to what extent *any* value in  $A$  is possible. When  $\pi$  reduces to  $E$ ,  $\Delta(A) = 1$  if  $A \subseteq E$ , and  $\Delta(A) = 0$  otherwise. Strong possibility measures are characterized by the following property:  $\Delta(A \cup B) = \min(\Delta(A), \Delta(B))$ .

<sup>1</sup> For instance, in [19] Lang et al. propose a formal theory of desires based on Boutilier’s logic QDT [4] in which two ordering relations representing preference and normality are given. The interpretation given to the statement “in context  $\varphi$ , I desire  $\psi$ ” is “the best among the most normal  $\varphi \wedge \psi$  worlds are preferred to the most normal  $\varphi \wedge \neg\psi$  worlds” which is different from interpretation of desire given in this paper.

- iv) The dual (*weak*) (or potential) necessity measure  $\nabla(A) = \max_{w \notin A} 1 - \pi(w) = 1 - \Delta(\bar{A})$  evaluates to what extent there is a value outside  $A$  that is impossible. When  $\pi$  reduces to  $E$ ,  $\nabla(A) = 1$  if  $A \cup E \neq U$ , and  $\nabla(A) = 0$  otherwise. Weak necessity measures are characterized by property:  $\nabla(A \cap B) = \max(\nabla(A), \nabla(B))$ .

$\Delta$ ,  $\nabla$  are decreasing set functions, while the (weak) possibility and (strong) necessity measures are increasing. A modal logic counterpart of these 4 modalities has been proposed in the *binary*-valued case (things are possible or impossible) [7]. There is a close link between Spohn functions and (weak) possibility / (strong) necessity measures [9].

### 3 Possibility Theory as Basis for a Logical Theory of Desires

The possibility and necessity operators  $\Pi$  and  $N$  have a clear epistemic meaning both in the frameworks of possibility theory, and of Spohn's uncertainty theory [27] (also referred to as ' $\kappa$  calculus', or as 'rank-based system' and 'qualitative probabilities' [16]). Differently from the operators  $\Pi$  and  $N$ , the operators  $\Delta$  and  $\nabla$  have not an intuitive interpretation in terms of epistemic attitudes. Indeed, although  $\Delta$  and  $\nabla$  make sense from the point of view of possibility theory and also from a logical viewpoint, it is not fully clear which kind of mental attitudes these two operators aim at modeling. Here we defend the idea that  $\Delta$  and  $\nabla$  can be viewed as operators modeling motivational mental attitudes such as goals or desires.<sup>2</sup> In particular, we claim that  $\Delta$  can be used to model the notion of *desire*, whereas  $\nabla$  can be used to model the notion of *potential* desire.<sup>3</sup>

According to the philosophical theory of motivation based on Hume [18], a desire can be conceived as an agent's motivational attitude which consists in an anticipatory mental representation of a pleasant (or desirable) state of affairs (representational dimension of desires) that motivates the agent to achieve it (motivational dimension of desires). In this perspective, the motivational dimension of an agent's desire is realized through its representational dimension. For example when an agent desires to be at the Japanese restaurant eating sushi, he imagines himself eating sushi at the Japanese restaurant and this representation gives him pleasure. This pleasant representation motivates him to go to the Japanese restaurant in order to eat sushi.

Intuitively speaking, with the term *potential* desire, we refer to a weaker form of motivational attitude. We assume that an agent considers a given property  $\varphi$  potentially desirable if  $\varphi$  does not conflict with the agent's current desires. In this sense,  $\varphi$  is potentially desirable if it is not incompatible with the agent's current desires. Following ideas presented in [21], let us explain why the operator  $\Delta$  is a good candidate for modeling the concept of desire and why  $\nabla$  is a good candidate for modeling the idea of desire compatibility. We define an agent's *mental state* as a tuple  $M = (E, D)$  where:

- $E \subseteq W$  is a *non-empty* subset of the set of all worlds, and
- $D \subset W$  is a proper subset of the set of all worlds.

<sup>2</sup> We use the term 'motivational' mental attitude (e.g., a desire, a goal or an intention) in order to distinguish it from an 'epistemic' mental attitude such as knowledge or belief.

<sup>3</sup> Here, the word *potential* does not refer to the idea that  $\varphi$  would be desired by the agent as a consequence of his mental state, but the agent has not enough deductive power to become aware of it. It is more the idea that the agent has no reason not to desire  $\varphi$ . Another possible term is *desire admissibility* or *desire compatibility*.

The set  $E$  defines the set of worlds envisaged by the agent (i.e., the set of worlds that the agent considers possible), whereas  $D$  is the set of desirable worlds for the agent. Let  $\mathcal{M}$  denote the set of all mental states. We here assume for every mental state  $M$  there exists a world with a minimal degree of desirability 0 (this is why  $D \neq W$ ). This type of normality constraint for guaranteed possibility distributions is usually assumed in possibility theory. More generally, a *graded mental state* is a pair  $M = (\pi, \delta)$  where:

- $\pi : W \rightarrow L$  is a normal possibility distribution over the set of all worlds, where ‘normal’ means that  $\pi(w) = 1$  for some  $w \in W$ , and
- $\delta : W \rightarrow L$  is a function mapping every world  $w$  to its desirability (or pleasantness) degree in  $L$ , with  $\delta(w) = 0$  for some  $w \in W$ .
- $L$  is a bounded chain acting as a qualitative scale for possibility and desirability, that make these notions commensurate.

Note that while  $\delta(w) = 1$  expresses complete desirability,  $\delta(w) = 0$  expresses indifference, rather than repulsion. The condition  $\delta(w) = 0$  for some  $w \in W$  indicates that desire presupposes that not everything is desired.

### 3.1 Modeling Desire Using $\Delta$ Function

We here assume that in order to determine how much a proposition  $\varphi$  is desirable an agent takes into consideration the worst situation in which  $\varphi$  is true. Thus, denoting by  $\|\varphi\|$  the set of situations where  $\varphi$  is true, for all graded mental states  $M = (\pi, \delta)$  and for all propositions  $\varphi$ , we can interpret  $\Delta(\|\varphi\|) = \min_{u \in \|\varphi\|} D(u)$  as the extent to which the agent desires  $\varphi$  to be true. Let us justify the following two properties for desires:  $\Delta(\|\varphi \vee \psi\|) = \min(\Delta(\|\varphi\|), \Delta(\|\psi\|))$  and  $\Delta(\|\varphi \wedge \psi\|) \geq \max(\Delta(\|\varphi\|), \Delta(\|\psi\|))$ .

According to the first property, an agent desires  $\varphi$  to be true with a given strength  $\alpha$  and desires  $\psi$  to be true with a given strength  $\beta$  if and only if the agent desires  $\varphi$  or  $\psi$  to be true with strength equal to  $\min(\alpha, \beta)$ . Notice that in the case of epistemic states, this property would not make any sense because the plausibility of  $\varphi \vee \psi$  should be clearly *at least* equal to the maximum of the plausibilities of  $\varphi$  and  $\psi$ . For the notion of desires, it seems intuitively satisfactory to have the opposite, namely the level of desire of  $\varphi \vee \psi$  should be *at most* equal to the minimum of the desire levels of  $\varphi$  and  $\psi$ . Indeed, we only deal with here with “*positive*”<sup>4</sup> desires (i.e., desires to reach something with a given strength). Under this proviso, the level of desire of  $\varphi \wedge \psi$  cannot be less than the maximum of the levels of desire of  $\varphi$  and  $\psi$ . According to the second property, the joint occurrence of two desired events  $\varphi$  and  $\psi$  is more desirable than the occurrence of one of the two events. This is the reason why in the right side of the equality we have the  $\max$ . The latter property does not make any sense in the case of epistemic attitudes like beliefs, as the joint occurrence of two events  $\varphi$  and  $\psi$  is epistemically less plausible than the occurrence of a single event. On the contrary it makes perfect sense for motivational attitudes like desires. By way of example, suppose Peter wishes to go to the cinema in the evening with strength  $\alpha$  (i.e.,  $\Delta(\|goToCinema\|) = \alpha$ )

<sup>4</sup> The distinction between positive and negative desires is a classical one in psychology. Negative desires correspond to state of affairs the agent wants to avoid with a given strength, and then desires the opposite to be true. However, we do not develop this bipolar view here.

and, at the same time, he wishes to spend the evening with his girlfriend with strength  $\beta$  (i.e.,  $\Delta(||\text{stayWithGirlfriend}||) = \beta$ ). Then, according to the preceding property, Peter wishes to go to the cinema with his girlfriend with strength at least  $\max\{\alpha, \beta\}$  (i.e.,  $\Delta(||\text{goToCinema} \wedge \text{stayWithGirlfriend}||) \geq \max\{\alpha, \beta\}$ ). This is a reasonable conclusion because the situation in which Peter achieves his two desires is (for Peter) at least as pleasant as the situation in which he achieves only one desire. A similar intuition can be found in [5] about the min-decomposability of disjunctive desires, where however it is emphasized that it corresponds to a pessimistic view.

From the normality constraint of  $\delta$ , we can deduce the following inference rule:

**Proposition 1.** *For every  $M \in \mathcal{M}$ , if  $\Delta(||\varphi||) > 0$  then  $\Delta(||\neg\varphi||) = 0$ .*

This means that if an agent desires  $\varphi$  to be true — i.e., with some strength  $\alpha > 0$  — then he does not desire  $\varphi$  to be false. In other words, an agent's desires must be consistent.

Note that the operator  $\Delta$  satisfies the following additional property:

**Proposition 2.** *For every  $M \in \mathcal{M}$ , if  $||\varphi|| = \emptyset$  then  $\Delta(||\varphi||) = 1$ .*

i.e., in absence of actual situations where  $\varphi$  is true, the property  $\varphi$  is desirable by default.

### 3.2 Modeling Potential Desire Using $\nabla$

As pointed out above, we claim that the operator  $\nabla$  allows us to capture a concept of potential desire (or desire compatibility):  $\nabla(||\varphi||)$  represents the extent to which an agent considers  $\varphi$  a potentially desirable property or, alternatively, the extent to which the property  $\varphi$  is not incompatible with the agent's desires. An interesting situation is when the property  $\varphi$  is *maximally* potentially desirable for the agent (i.e.,  $\nabla(||\varphi||) = 1$ ). This is the same thing as saying that the agent does not desire  $\varphi$  to be false (i.e.,  $\Delta(||\neg\varphi||) = 0$ ). Intuitively, this means that  $\varphi$  is totally potentially desirable in as much as the level of desire for  $\neg\varphi$  is 0. In particular, given a graded mental state  $M = (\pi, \delta)$ , let  $D = \{w \in W : \delta(w) > 0\}$  be the set of somewhat satisfactory or desirable worlds in  $M$ . Then, we have  $\nabla(||\varphi||) = 1$  if and only if  $\overline{D} \cap ||\neg\varphi|| \neq \emptyset$ , i.e.,  $\neg\varphi$  is consistent with what is not desirable, represented by the set  $\overline{D}$ .

Another interesting situation is when the property  $\varphi$  is *maximally* desirable for the agent (i.e.,  $\Delta(||\varphi||) = 1$ ). This is the same thing as saying that  $\neg\varphi$  is not at all potentially desirable for the agent (i.e.,  $\nabla(||\neg\varphi||) = 0$ ). It is worth noting that if an agent desires  $\varphi$  to be true, then  $\varphi$  should be *maximally* potentially desirable. This property is expressed by the following valid inference rule which follows straightforwardly from the previous one and from the definition of  $\nabla(||\varphi||)$  as  $1 - \Delta(||\neg\varphi||)$ :

**Proposition 3.** *For every  $M$ , if  $\Delta(||\varphi||) > 0$  then  $\nabla(||\varphi||) = 1$ .*

Let us now consider the case in which the agent does not desire  $\varphi$  (i.e.,  $\Delta(||\varphi||) = 0$ ). In this case two different situations are possible: either  $\Delta(||\neg\varphi||) = 0$  and  $\varphi$  is *fully* compatible with the agent's desires (i.e.,  $\nabla(||\varphi||) = 1$ ), or  $\Delta(||\neg\varphi||) > 0$  and then  $\varphi$  is not *fully* compatible with the agent's desires (i.e.,  $\nabla(||\varphi||) < 1$ ).

### 3.3 Some Valid Inference Rules for Desires

The following is a valid inference rule for  $\Delta$ -based logic, see [7,12] for the proof:



**Proposition 4.** For every  $M \in \mathcal{M}$ , if  $\Delta(\|\varphi \wedge \psi\|) \geq \alpha$  and  $\Delta(\|\neg\varphi \wedge \chi\|) \geq \beta$  then  $\Delta(\|\psi \wedge \chi\|) \geq \min(\alpha, \beta)$ .

Therefore, if we interpret  $\Delta$  as a desire operator, we have that if an agent desires  $\varphi \wedge \psi$  with strength at least  $\alpha$  and desires  $\neg\varphi \wedge \chi$  with strength at least  $\beta$ , then he desires  $\psi \wedge \chi$  with strength at least  $\min(\alpha, \beta)$ . This seems a reasonable property of desires. By way of example, suppose Peter desires to be in a situation in which he drinks red wine and eats a pizza with strength at least  $\alpha$  and, at the same time, he desires to be in a situation in which he does not drink red wine and eats tiramisú as a dessert with strength at least  $\beta$ . Then, it is reasonable to conclude that Peter desires to be in a situation in which he eats both a pizza and tiramisú with strength at least  $\min(\alpha, \beta)$ .

Another rule, never studied, mixes  $\Delta$  (*alias* actual desire) and  $\nabla$  (potential desire):

**Proposition 5.** For every  $M \in \mathcal{M}$ , if  $\Delta(\|\varphi \wedge \psi\|) \geq \alpha$  and  $\nabla(\|\neg\varphi \wedge \chi\|) \geq \beta$  then  $\nabla(\|\psi \wedge \chi\|) \geq \alpha * \beta$ , where  $\alpha * \beta = \alpha$  if  $\alpha > 1 - \beta$  and  $\alpha * \beta = 0$  if  $1 - \beta \geq \alpha$ .

*Proof.* First, we have by duality  $\Delta(\|\varphi \wedge \psi\|) \geq \alpha \Leftrightarrow \nabla(\|\neg\varphi \vee \neg\psi\|) \leq 1 - \alpha$ . Then observe  $\neg\varphi \wedge \chi \equiv (\neg\varphi \vee \neg\psi) \wedge (\neg\varphi \vee \psi) \wedge \chi$ . Thus  $\nabla(\|\neg\varphi \wedge \chi\|) = \max(\nabla(\|\neg\varphi \vee \neg\psi\|), \nabla(\|(\neg\varphi \vee \psi) \wedge \chi\|)) \geq \beta$  which leads to  $\max(1 - \alpha, \nabla(\|\psi \wedge \chi\|)) \geq \beta$  from which the result follows. The last inequality is obtained by noticing that  $\nabla(\|(\neg\varphi \vee \psi) \wedge \chi\|) \leq \nabla(\|\psi \wedge \chi\|)$  due to the decreasingness of  $\nabla$ . It can be shown that  $\alpha * \beta$  is the tightest lower bound that can be established for the above pattern.  $\square$

Thus, in particular, if  $\varphi$  is *fully* potentially desirable ( $\nabla(\|\varphi\|) = 1$ ), and  $\neg\varphi \wedge \psi$  is *fully* desirable ( $\Delta(\|\neg\varphi \wedge \psi\|) = 1$ ), then  $\psi$  is *fully* potentially desirable ( $\nabla(\|\psi\|) = 1$ ). The two above inference rules are the counterparts of the following inference rule:

if  $N(\|\varphi \vee \psi\|) \geq \alpha$  and  $N(\|\neg\varphi \vee \chi\|) \geq \beta$  then  $N(\|\psi \vee \chi\|) \geq \min(\alpha, \beta)$   
(the basic inference rule in standard possibilistic logic), and of the following one [8]:

if  $N(\|\varphi \vee \psi\|) \geq \alpha$  and  $\Pi(\|\neg\varphi \vee \chi\|) \geq \beta$  then  $\Pi(\|\psi \vee \chi\|) \geq \alpha * \beta$

with  $\alpha * \beta = \alpha$  if  $\alpha > 1 - \beta$  and  $\alpha * \beta = 0$  if  $1 - \beta \geq \alpha$ . They are themselves the graded counterparts of two inference rules well-known in modal logic [13,8].

### 3.4 Realistic Desires

Besides,  $\Delta(\|\varphi\|) = \alpha$  implies that for any  $\psi$  logically independent from  $\varphi$ , it holds that  $\Delta(\|\varphi \wedge \psi\|) \geq \alpha$  and  $\Delta(\|\varphi \wedge \neg\psi\|) \geq \alpha$ , which may sound counterintuitive. Indeed, suppose you wish to choose a menu and you prefer to eat fish than not to a certain degree. It means that you should wish to eat fish with white wine, and fish with red wine to a degree at least as high. Yet, you may dislike very much to drink red wine with fish. Your desire for fish presupposes the restaurant offers white wine as well. So you express your desire for fish is conditioned to the possibility of having white wine as well. In other words, you believe that in fish restaurants it is more likely to find white wine than red wine. Modeling desire irrespectively of what you assume to be possible is liable to such kind of paradoxes when using the set-function  $\delta$ .

This discussion suggests that a *realistic desire* can be defined as one whose realization is considered possible by the agent. The most natural representation consists



in restricting mental states to pairs  $(E, D)$  such that  $E \cap D \neq \emptyset$ , called *realistic mental states*, the set of which is denoted by  $\mathcal{M}^r$ . Indeed, if  $E \cap D = \emptyset$ , then the agent knows that desirable states are impossible in his view. Then the agent with mental state  $(E, D)$  is said to *realistically desire*  $\varphi$  if and only if  $\Delta(\|\varphi\|) = 1$  (that is,  $\|\varphi\| \subseteq D$ ) and  $\Pi(\|\varphi\|) = 1$  (that is,  $\|\varphi\| \cap E \neq \emptyset$ ).

A more conservative notion of realistic desire would consist in requesting  $N(\|\varphi\|) = 1$  instead of  $\Pi(\|\varphi\|) = 1$ , that is realistic desire would concern only propositions  $\varphi$  such that the agent is certain that  $\varphi$  is true. However one may question the fact that realistically desiring  $\varphi$  to be true may presuppose no risk at all for  $\varphi$  being false, namely the complete certainty that  $\varphi$  is true. This corresponds better to the idea of happiness. On the contrary, the preceding notion of realistic desire defined by  $\Delta(\|\varphi\|) = 1$  and  $\Pi(\|\varphi\|) = 1$  corresponds to the notion of hope. Indeed, according to psychological theories of emotion (e.g., [24]), while happiness is triggered by *prospective consequences* (or *prospects*), hope is triggered by *actual consequences*.<sup>5</sup>

In the case of graded mental states,  $(\pi, \delta)$ , one may take a restrictive point of view on possible states of affairs, evaluating desired statements over pairs  $(\pi, \delta)$  such that  $E = \{w : \pi(w) = 1\} \cap \{w : \delta(w) > 0\} \neq \emptyset$ . It comes down to working with pairs  $(E, \delta)$  where only desire is graded. Then desires are expressed under the assumption that they can be achieved in at least one normal situation.

Alternatively one may compute the degree of realistic desire in the mental state  $(\pi, \delta)$  as  $\rho(\|\varphi\|) = \min(\Delta(\|\varphi\|), \Pi(\|\varphi\|))$ . It presupposes that degrees of plausibility and degrees of desire are commensurate.

Note that the above proposal differs from the one that would restrict desired states to possible ones, that is, replacing  $\delta$  by  $\delta_\pi = \min(\delta, \pi)$  since  $\rho(A) \geq \Delta_\pi(A) = \min_{w \in A} \min(\delta(w), \pi(w))$ . For instance, if  $E \cap D \neq \emptyset$ ,  $A \cap E \neq \emptyset$ ,  $\overline{E} \cap A \neq \emptyset$ , and  $A \subseteq D$ , then  $\rho(A) = 1$  but  $\Delta_\pi(A) = 0$  since then  $\delta_\pi(w) = 0$  for some  $w \in A$ .

## 4 Logics of Beliefs and Desires

In this section we introduce some variants of a modal logic of beliefs and desires, called here BDL, based on the ideas presented in the previous sections. Specifically, the logics presented here support reasoning about the notion of belief, as traditionally studied in the area of modal logic of belief (*alias* doxastic logic) [23,14,17], in combination with the notion of ( $\Delta$ -based) desire discussed in Section 3.1 and the notion of ( $\Delta$ -based) realistic desire discussed in Section 3.4. We first consider a simpler logic, in line with the previous sections, that, like MEL [1], does not support the nesting of modalities and allows us to reason about purely (non-graded) notions of belief and desire. The semantics will then be defined in terms of mental states (pairs of sets or distributions). Then, we present a simple generalization of this logic that allows us to formalize notions of graded belief and graded desire. Finally, we consider a full-fledged modal logic of graded beliefs and graded desires with multiple agents that supports the nesting of modalities. The nesting of modalities is crucial in order to represent an agent  $i$ 's beliefs about the beliefs (or the desires) of a different agent  $j$ .

<sup>5</sup> Like [15], we here interpret the term ‘prospect’ as synonymous of ‘uncertain consequence’ (in contrast with ‘actual consequence’ as synonymous of ‘certain consequence’).

#### 4.1 Minimal Modal Logic of Beliefs and Desires MBDL

Let us introduce a propositional language  $PL$  based on a countable set  $Prop$  of atomic propositions (with typical members denoted  $p, q, \dots$ ), and defined by the following grammar:  $\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \psi$ , where  $p$  ranges over a given countable set of atomic propositions  $Prop = \{p, q, \dots\}$ , some of which can be decision variables. The other Boolean constructions  $\top, \perp, \vee, \rightarrow$  and  $\leftrightarrow$  are defined from  $p, \neg$  and  $\wedge$  in the standard way. A propositional valuation is defined in the standard way as a subset of atomic propositions considered as true, the other ones being false. Propositional valuations, also called *worlds* or *states*, are denoted by symbols  $w$ . The set  $W$  is identified with the set  $2^{Prop}$  of all propositional valuations. Let  $\|p\| = \{w : p \in w\}$  be the set of models of  $p$ .

The extension of propositional formulas is defined in the standard way as follows:

$$\|\neg\varphi\| = W \setminus \|\varphi\|; \|\varphi \wedge \psi\| = \|\varphi\| \cap \|\psi\|$$

We first consider the Boolean case and the most elementary language that may capture the previously introduced notions. The language  $\mathcal{L}_{N,\Delta}$  of the logic MBDL is defined as follows:  $\Phi ::= N\varphi \mid \Delta\varphi \mid \neg\Phi \mid \Phi \wedge \Psi$ , where formulas  $\varphi$  range over  $PL$ . In other words,  $N\varphi$  and  $\Delta\varphi$  are atomic propositions, respectively referring to the statements “the agent believes that  $\varphi$ ” and “the agent desires that  $\varphi$ ”. The two modal operators  $N$  and  $\Delta$  have the following intuitive readings:

- $N\varphi$  means: the agent believes that  $\varphi$  is true (i.e.,  $\varphi$  is true in all worlds that the agent envisages as possible),
- $\Delta\varphi$  means: the agent considers  $\varphi$  desirable in all worlds where  $\varphi$  is true.

The dual operators  $\Pi$  and  $\nabla$  are defined in the usual way as follows:

$$\Pi\varphi \stackrel{\text{def}}{=} \neg N\neg\varphi; \quad \nabla\varphi \stackrel{\text{def}}{=} \neg\Delta\neg\varphi$$

The set of axioms of MBDL is given in Figure 2 and provides a proof system for this logic MBDL. The first three modal axioms are those of KD, more specifically its subjective fragment where modalities are not nested. They account for Boolean necessity measures. The three following ones for the desire modality are the translation of the former when replacing necessity by guaranteed possibility, using the identity  $N(A) = \min_{w \in A} 1 - \pi(w) = \Delta(\bar{A})$  if  $\delta = 1 - \pi$ .

All tautologies of propositional calculus	(PC)
$(N\varphi \wedge N(\varphi \rightarrow \psi)) \rightarrow N\psi$	( $\mathbf{K}_N$ )
$\neg(N\varphi \wedge N\neg\varphi)$	( $\mathbf{D}_N$ )
$N\top$	( $\mathbf{N}_N$ )
$(\Delta\varphi \wedge \Delta(\neg\varphi \wedge \psi)) \rightarrow \Delta\psi$	( $\mathbf{K}_\Delta$ )
$\neg(\Delta\varphi \wedge \Delta\neg\varphi)$	( $\mathbf{D}_\Delta$ )
$\Delta\perp$	( $\mathbf{N}_\Delta$ )
$\frac{\varphi, \varphi \rightarrow \psi}{\psi}$	(MP)

**Fig. 1.** Sound and complete axiomatization of MBDL

The truth of a MBDL formula is evaluated w. r. t. a valuation  $w$  and a mental state  $M = (E, D) \in \mathcal{M}$ , by means of the following rules:

$$\begin{aligned} M \models N\varphi &\iff \forall w \in E, w \in \|\varphi\|; & M \models \Delta\varphi &\iff \forall w \models \varphi, w \in D \\ M \models \neg\Phi &\iff M \not\models \Phi; & M \models \Phi \wedge \Psi &\iff M \models \Phi \text{ AND } M \models \Psi \end{aligned}$$

We say that a formula  $\Phi$  of the language  $\mathcal{L}_{\text{BDL}}(\text{Prop})$  is valid, denoted by  $\models_{\text{MBDL}} \Phi$ , if and only if for every mental state  $M$  in  $\mathcal{M}$   $M \models \Phi$ . We say that  $\Phi$  is satisfiable if and only if  $\neg\Phi$  is not valid. It can be checked that

- $M \models N\varphi$  if and only if  $N(\|\varphi\|) = 1$  with respect to  $E$ .
- $M \models \Delta\varphi$  if and only if  $\Delta(\|\varphi\|) = 1$  with respect to  $D$ .

In the logic MBDL we can also formally express the concept of realistic desire:

$$R\Delta\varphi \stackrel{\text{def}}{=} \Delta\varphi \wedge \neg N\neg\varphi.$$

We can now prove the completeness theorem for this logic:

**Theorem 1.** *The axioms and the rules of inference given in Figure 1 provides a sound and complete axiomatization for the logic MBDL.*

*Proof (Sketch).* Soundness is easy to obtain. As to completeness, note that axioms  $K_N$ ,  $D_N$ , and  $N_N$  imply the equivalence between  $N(\varphi \wedge \psi)$  and  $N\varphi \wedge N\psi$ . Likewise, axioms  $K_\Delta$ ,  $D_\Delta$ , and  $N_\Delta$  imply the equivalence between  $\Delta(\varphi \vee \psi)$  and  $\Delta\varphi \wedge \Delta\psi$ . Besides, a propositional valuation  $v$  of the language of MBDL assigns 0 or 1 to each  $N\varphi$  and  $\Delta\varphi$ . Define two set functions  $\nu$  and  $\mu$  over  $W$  by letting  $\nu(A) = 1$  if and only if  $v(N\varphi) = 1$  if  $A = \|\varphi\|$  and 0 otherwise; and likewise  $\mu(A) = 1$  if and only if  $v(\Delta\varphi) = 1$  if  $A = \|\varphi\|$  and 0 otherwise. Axioms of propositional logic ensures these definitions are sound (truth assignments to  $N\varphi$  and  $\Delta\varphi$  do not change if  $\varphi$  is replaced by a logically equivalent proposition). Moreover the 6 first modal axioms imply that  $\nu$  is a Boolean necessity measure and  $\mu$  a Boolean guaranteed possibility measure. It means that there exists a mental state  $M = (E_v, D_v)$  such that  $v(N\varphi) = 1$  if and only if  $E_v \subseteq \|\varphi\|$  and  $v(\Delta\varphi) = 1$  if and only if  $\|\varphi\| \subseteq D_v$ . Using the completeness of propositional logic, we thus prove that, for any subset of formulas  $B$  in the logic MBDL:

$$\begin{aligned} B \vdash_{\text{MBDL}} \Phi &\iff B \cup \{K_N, D_N, N_N, K_\Delta, D_\Delta, N_\Delta\} \vdash_{PL} \Phi \\ &\iff (\forall v, v \models_{PL} B \cup \{K_N, D_N, N_N, K_\Delta, D_\Delta, N_\Delta\} \Rightarrow v \models_{PL} \Phi) \\ &\iff \forall (E, D) \in \mathcal{M}, (E, D) \models_{\text{MBDL}} B \Rightarrow (E, D) \models_{\text{MBDL}} \Phi \iff B \models_{\text{MBDL}} \Phi \end{aligned}$$

An interesting aspect of the modal logic MBDL is that one can make syntactic proofs of some properties of the notions of ( $\Delta$ -based) desire. For instance, we can give the following syntactic proof of the inference rule for desire given in Proposition 4 when  $\alpha = 1$ , namely  $\{\Delta(\varphi \wedge \psi), \Delta(\neg\varphi \wedge \chi)\} \vdash_{\text{MBDL}} \Delta(\psi \wedge \chi)$ :

1. Applying  $K_\Delta$  when  $\psi \models \varphi$  yields theorem T1:  $\vdash_{\text{MBDL}} \Delta\varphi \rightarrow \Delta\psi$  if  $\psi \models \varphi$ .
2.  $K_\Delta$  can be written as T2: if  $\varphi \wedge \psi$  is a contradiction,  $\vdash_{\text{MBDL}} \Delta\varphi \wedge \Delta\psi \rightarrow \Delta(\varphi \vee \psi)$ .
3. Applying T2:  $\{\Delta(\varphi \wedge \psi), \Delta(\neg\varphi \wedge \chi)\} \vdash_{\text{MBDL}} \Delta((\varphi \wedge \psi) \vee (\neg\varphi \wedge \chi))$
4. By T1,  $\{\Delta(\varphi \wedge \psi), \Delta(\neg\varphi \wedge \chi)\} \vdash_{\text{MBDL}} \Delta(\psi \wedge \chi)$  since  $\psi \wedge \chi \models (\varphi \wedge \psi) \vee (\neg\varphi \wedge \chi)$

## 4.2 Outline of a Minimal Modal Logic of Graded Beliefs and Desires MGBDL

Assume a finite chain  $L \subseteq [0, 1]$  containing the values 0 and 1 such that for every  $\alpha \in L$  we have  $1 - \alpha \in L$ . For every  $\alpha \in L$  such that  $\alpha > 0$ , let  $p(\alpha)$  denote the number  $\beta \in L$  such that  $\beta < \alpha$  and there is no  $\gamma \in L$  such that  $\beta < \gamma < \alpha$ .  $\beta$  is called the predecessor of  $\alpha$  in  $L$ . Furthermore, let  $p(0) = 0$ . For every  $\alpha \in L$  such that  $\alpha < 1$ , let  $\sigma(\alpha)$  denote the number  $\beta \in L$  such that  $\alpha < \beta$  and there is no  $\gamma \in L$  such that  $\alpha < \gamma < \beta$ .  $\beta$  is called the successor of  $\alpha$  in  $L$ . Furthermore, let  $\sigma(1) = 1$ .

The language  $\mathcal{L}_{N,\Delta}^L$  of the graded logic MGBDL is defined as follows:

$$\Phi ::= N^{\geq \alpha} \varphi \mid \Delta^{\geq \beta} \varphi \mid \neg \Phi \mid \Phi \wedge \Psi,$$

where formulas  $\varphi$  range over  $PL$  and  $\alpha > 0, \beta > 0 \in L$ . In other words,  $N^{\geq \alpha} \varphi$  and  $\Delta^{\geq \beta} \varphi$  are atomic propositions, respectively encoding the statements:

- $N^{\geq \alpha} \varphi$  means: the agent believes that  $\varphi$  is true to at least level  $\alpha$  (i.e.,  $\varphi$  is true in all worlds that the agent envisages as possible at level at least  $\sigma(1 - \alpha)$ ),
- $\Delta^{\geq \beta} \varphi$ : for the agent  $\varphi$  is desirable at least at level  $\beta$  in all worlds where  $\varphi$  is true.

The set of axioms of MGBDL are those of PL, the six first modal ones of Figure 1 for each  $N^{\geq \alpha}$  and  $\Delta^{\geq \beta}$ . Finally, we must add weakening axioms for the two graded modalities:

$$[W_\Delta]: \Delta^{\geq \beta} \varphi \rightarrow \Delta^{\geq p(\beta)} \varphi \quad [W_N]: N^{\geq \alpha} \varphi \rightarrow N^{\geq p(\alpha)} \varphi.$$

The semantics is defined by means of graded mental states  $(\pi, \delta)$  as defined in Section 3. A completeness theorem can be proved as for the Boolean case. Again the idea is to interpret any propositional valuation  $v$  of the language  $\mathcal{L}_{N,\Delta}^L$  as a pair of set functions  $(g_N^v, g_\Delta^v)$  on  $W$  stemming from a pair  $(\pi, \delta)$ . Namely we can let  $g_N^v(|\varphi|) = \max\{\alpha : v(N^{\geq \alpha} \varphi) = 1\}$ , and  $g_\Delta^v(|\varphi|) = \max\{\beta : v(\Delta^{\geq \beta} \varphi) = 1\}$ , which is meaningful due to the weakening axioms, and prove that the other axioms ensure that  $g_N^v$  is a necessity measure, and  $g_\Delta^v$  is a guaranteed possibility measure.

## 4.3 Multi-agent Modal Logic of Graded Beliefs and Desires GBDL<sup>n</sup>

Let  $Agt = \{1, \dots, n\}$  be a finite set of agents (or individuals). The language  $\mathcal{L}_{N_i, \Delta_i}^L$  of the logic GBDL<sup>n</sup> consists of a set of formulae and is defined as follows:

$$\varphi ::= p \mid \neg \varphi \mid \varphi \wedge \psi \mid N_i^{\geq \alpha} \varphi \mid \Delta_i^{\geq \beta} \varphi$$

where  $p$  ranges over the set of atomic propositions  $Prop$ ,  $\alpha, \beta \in L \setminus \{0\}$  and  $i$  ranges over the set of agents  $Agt$ . The two modal operators  $N_i^{\geq \alpha}$  and  $\Delta_i^{\geq \beta}$  have the following intuitive readings:

- $N_i^{\geq \alpha} \varphi$  means: agent  $i$  believes that  $\varphi$  is true with strength at least  $\alpha$  (i.e.,  $\varphi$  is true in all worlds that agent  $i$  considers possible at level at least  $\sigma(1 - \alpha)$ ),
- $\Delta_i^{\geq \beta} \varphi$  means: agent  $i$  desires  $\varphi$  with strength at least  $\beta$  (i.e., all worlds in which  $\varphi$  is true are desirable for agent  $i$  at level at least  $\beta$ ).

The interesting aspect of the logic GBDL<sup>n</sup> is that it allows to represent what a given agent  $i$  believes about another  $j$ 's beliefs and desires. For instance, the formula  $N_1^{\geq \alpha_1} N_2^{\geq \alpha_2} \varphi$  expresses that agent 1 believes with strength at least  $\alpha_1$  that agent 2

believes  $\varphi$  with strength at least  $\alpha_2$ , whereas the formula  $N_1^{\geq \alpha_1} \Delta_2^{\geq \beta_2} \varphi$  expresses that agent 1 believes with strength at least  $\alpha_1$  that agent 2 desires  $\varphi$  with strength at least  $\beta_2$ .

The semantics of the logic GBDL<sup>n</sup> is defined in terms of multi-agent mental states of the form  $M = (S, \{\pi_{i,s}\}_{i \in \text{Agt}, s \in S}, \{\delta_{i,s}\}_{i \in \text{Agt}, s \in S}, V)$  where:

- $S$  is a set of states, including states of the agents (possibly more general than  $W$ ) ;
- for all  $s \in S$  and for all  $i \in \text{Agt}$ ,  $(\pi_{i,s}, \delta_{i,s})$  is a graded mental state over the set  $S$  as the one defined in Section 3;
- $V : S \rightarrow 2^{\text{Prop}}$  is a valuation function for atomic propositions:  $p \in V(s)$  means that proposition  $p$  is true at world  $w = V(s)$ .

Specifically,  $\pi_{i,s}(s')$  captures how much, in state  $s$ , agent  $i$  thinks that state  $s'$  is (epistemically) possible, while  $\delta_{i,s}(s')$  captures how much, in state  $s$ , agent  $i$  thinks that state  $s$  is desirable. Note that parameterizing possibility distributions  $\pi_{i,s}$  and  $\delta_{i,s}$  with states is one way to model an agent  $i$ 's uncertainty about the beliefs and the desires of another agent  $j$ .

The truth of a GBDL<sup>n</sup> formula is evaluated with respect to a given state  $s$  in a multi-agent mental state  $M = (S, \{\pi_{i,s}\}_{i \in \text{Agt}, s \in S}, \{\delta_{i,s}\}_{i \in \text{Agt}, s \in S}, V)$  by means of the following rules:

$$\begin{aligned}
M, s \models p &\iff p \in V(s) \\
M, s \models \neg \varphi &\iff M, s \not\models \varphi \\
M, s \models \varphi \wedge \psi &\iff M, s \models \varphi \text{ AND } M, s \models \psi \\
M, s \models N_i^{\geq \alpha} \varphi &\iff \forall s' \in S : \text{IF } \pi_{i,s}(s') \geq \sigma(1 - \alpha) \text{ THEN } M, s \models \varphi \\
M, s \models \Delta_i^{\geq \beta} \varphi &\iff \forall s' \in S : \text{IF } M, s \models \varphi \text{ THEN } \delta_{i,s}(s') \geq \beta
\end{aligned}$$

We say that a formula  $\varphi$  of the logic GBDL<sup>n</sup> is valid, denoted by  $\models_{\text{GBDL}^n} \varphi$ , if and only if for every multi-agent mental state  $M = (S, \{\pi_{i,s}\}_{i \in \text{Agt}, s \in S}, \{\delta_{i,s}\}_{i \in \text{Agt}, s \in S}, V)$  and for every state  $s$  in  $S$ ,  $M, s \models \varphi$ . We say that  $\varphi$  is GBDL<sup>n</sup> satisfiable if and only if  $\neg \varphi$  is not GBDL<sup>n</sup> valid. For instance,  $\neg N_1^{\geq \sigma(0)} \Delta_2^{\geq \sigma(0)} p \wedge \neg N_1^{\geq \sigma(0)} \neg \Delta_2^{\geq \sigma(0)} p$  is a satisfiable formula in the logic GBDL<sup>n</sup>. It means that agent 1 is uncertain whether agent 2 desires  $p$  or not.

In the logic GBDL<sup>n</sup> we can also formally express the concept of realistic desire for a given agent  $i$ . For all  $i \in \text{Agt}$  we define:

$$R\Delta_i^{\geq \beta} \varphi \stackrel{\text{def}}{=} \Delta_i^{\geq \beta} \varphi \wedge \neg N_i^{\geq 1} \neg \varphi$$

where  $R\Delta_i^{\geq \beta} \varphi$  has to be read “agent  $i$  *realistically* desires  $\varphi$  with strength at least  $\beta$ ”. The above realistic desire operator  $R\Delta_i^{\geq \beta}$  exactly corresponds to the notion of realistic desire discussed in Section 3.4. The general idea is that agent  $i$  *realistically* desires  $\varphi$  with strength at least  $\beta$  (i.e.,  $R\Delta_i^{\geq \beta} \varphi$ ) if and only if agent  $i$  desires  $\varphi$  with strength at least  $\beta$  and is not completely certain that  $\varphi$  is false.

We can prove that the list of principles given in Figure 2 provides a proof system for the logic GBDL<sup>n</sup>. In the axiomatization we use the following abbreviation for all  $\gamma \in L \setminus \{1\}$ :

$$\Box_i^{\leq \gamma} \varphi \stackrel{\text{def}}{=} \Delta_i^{\geq \sigma(\gamma)} \neg \varphi$$

All tautologies of propositional calculus	(PC)
$(N_i^{\geq \alpha} \varphi \wedge N_i^{\geq \alpha} (\varphi \rightarrow \psi)) \rightarrow N_i^{\geq \alpha} \psi$	( <b>K</b> $_{N_i^{\geq \alpha}}$ )
$\neg(N_i^{\geq \alpha} \varphi \wedge N_i^{\geq \alpha} \neg \varphi)$	( <b>D</b> $_{N_i^{\geq \alpha}}$ )
$(\Box_i^{\leq \gamma} \varphi \wedge \Box_i^{\leq \gamma} (\varphi \rightarrow \psi)) \rightarrow \Box_i^{\leq \gamma} \psi$	( <b>K</b> $_{\Box_i^{\leq \gamma}}$ )
$\neg(\Box_i^{\leq 0} \varphi \wedge \Box_i^{\leq 0} \neg \varphi)$	( <b>D</b> $_{\Box_i^{\leq 0}}$ )
$N_i^{\geq \alpha} \varphi \rightarrow N_i^{\geq p(\alpha)} \varphi$	( <b>Incl</b> $_{N_i^{\geq \alpha}, N_i^{\geq p(\alpha)}}$ )
$\Box_i^{\leq \gamma} \varphi \rightarrow \Box_i^{\leq p(\gamma)} \varphi$	( <b>Incl</b> $_{\Box_i^{\leq \gamma}, \Box_i^{\leq p(\gamma)}}$ )
$\frac{\varphi, \varphi \rightarrow \psi}{\psi}$	(MP)
$\frac{\varphi}{N_i^{\geq \alpha} \varphi}$	( <b>Nec</b> $_{N_i^{\geq \alpha}}$ )
$\frac{\varphi}{\Box_i^{\leq \gamma} \varphi}$	( <b>Nec</b> $_{\Box_i^{\leq \gamma}}$ )

**Fig. 2.** Sound and complete axiomatization of  $\text{GBDL}^n$

**Theorem 2.** *The axioms and the rules of inference given in Figure 2 provides a sound and complete axiomatization for the logic  $\text{BDL}$ .*

*Proof (Sketch).* It is a routine task to verify that the axioms given in Figure 2 are sound and that the rules of inference preserve validity. The proof of completeness has 2 steps. Step 1 consists in proving that the semantics of the logic  $\text{GBDL}^n$  given above is equivalent to an alternative semantics in terms of Kripke models with accessibility relations. Specifically, let us define the notion of Kripke  $\text{GBDL}^n$  model as a tuple  $M = (S, \{T_{i, \geq \alpha}\}_{i \in \text{Agt}, \alpha \in L \setminus \{0\}}, \{R_{i, \leq \gamma}\}_{i \in \text{Agt}, \gamma \in L \setminus \{1\}}, V)$  where  $S$  and  $V$  are as defined above, and every  $T_{i, \geq \alpha}$  and every  $R_{i, \leq \gamma}$  are binary relations on  $S$  satisfying the constraints: (C1)  $T_{i, \geq 1}$  is serial; (C2)  $R_{i, \leq 0}$  is serial; (C3) for all  $\alpha \in L \setminus \{0\}$ ,  $T_{i, \geq \sigma(\alpha)} \subseteq T_{i, \geq \alpha}$ ; (C4) for all  $\gamma \in L \setminus \{1\}$ ,  $R_{i, \leq p(\gamma)} \subseteq R_{i, \leq \gamma}$ .

In this alternative Kripke semantics for  $\text{GBDL}^n$ , the truth of a formula is evaluated w. r. t. a state  $s$  in a Kripke  $\text{GBDL}^n$  model  $M$  by means of the following rules:

$$\begin{aligned}
M, s \models p &\iff p \in V(s) ; M, s \models \neg \varphi \iff M, s \not\models \varphi \\
M, s \models \varphi \wedge \psi &\iff M, s \models \varphi \text{ AND } M, s \models \psi \\
M, s \models N_i^{\geq \alpha} \varphi &\iff \forall s' \in T_{i, \geq \sigma(1-\alpha)}(s) : M, s' \models \varphi \\
M, s \models \Delta_i^{\geq \beta} \varphi &\iff \forall s' \in R_{i, \leq p(\beta)}(s) : M, s' \models \neg \varphi
\end{aligned}$$

where  $T_{i, \geq \sigma(1-\alpha)}(s) = \{s' \in S \mid (s, s') \in T_{i, \geq \sigma(1-\alpha)}\}$  and  $R_{i, \leq p(\beta)}(s) = \{s \in S \mid (s, s') \in R_{i, \leq p(\beta)}\}$ . We say that a formula  $\varphi$  of the logic  $\text{GBDL}^n$  is valid with respect to the class of Kripke  $\text{GBDL}^n$  models if and only if for every Kripke  $\text{GBDL}^n$  model  $M$  and for every state  $s$  in  $S$  we have  $M, s \models \varphi$ .

**Lemma 1.** *For every formula  $\varphi$  of the logic  $\text{GBDL}^n$ ,  $\models_{\text{GBDL}^n} \varphi$  if and only if  $\varphi$  is valid with respect to the class of Kripke  $\text{GBDL}^n$  models.*

The 2nd step of the proof consists in showing that the list of principles given in Figure 2 completely axiomatizes the set of validities of the logic  $\text{GBDL}^n$  whose language is interpreted over Kripke  $\text{GBDL}^n$  models. It is a routine task to check that the axioms in Figure 2 correspond one-to-one to their semantic counterparts on  $\text{GBDL}^n$  Kripke models. In particular, Axioms  $\mathbf{K}_{N_i^{\geq \alpha}}$  and  $\mathbf{K}_{\Box_i^{\leq \gamma}}$  together with the rules of inference  $\mathbf{Nec}_{N_i^{\geq \alpha}}$  and  $\mathbf{Nec}_{\Box_i^{\leq \gamma}}$  correspond to the fact that  $N_i^{\geq \alpha}$  and  $\Box_i^{\leq \gamma} \varphi$  are normal modal operators interpreted by means of accessibility relations. Axiom  $\mathbf{D}_{N_i^{\geq \alpha}}$  corresponds to the fact that the relation  $T_{i, \geq 1}$  is serial (Constraint C1), while Axiom  $\mathbf{D}_{\Box_i^{\leq 0}}$  corresponds to the fact that the relation  $R_{i, \leq 0}$  is serial (Constraint C2). Moreover, Axioms  $\mathbf{Incl}_{N_i^{\geq \alpha}, N_i^{\geq p(\alpha)}}$  and  $\mathbf{Incl}_{\Box_i^{\leq \gamma}, \Box_i^{\leq p(\gamma)}}$  correspond respectively to the Constraints C3 and C4.

It is routine, too, to check that all principles given in Figure 2 are in the so-called Sahlqvist class [25]. This means that they are complete with respect to the defined model classes, cf. [3, Th. 2.42].  $\square$

## 5 Conclusive Remarks: Towards Emotions

In the previous sections, we have shown that possibility theory offers a unified logical framework in which both epistemic attitudes such as beliefs and motivational attitudes such as desires can be modeled. As a perspective along this line, we may study how the components of the approach, the epistemic one and the motivational one, can be combined in order to model basic emotion types such as hope and fear. Similar ideas on the logic of emotion intensity have been recently presented in [6] without making a connection with possibility theory.

Besides, we have described two extreme approaches to the problem: a minimal single agent logic of desire and belief and a maximal multi-agent one. The first one is completely faithful to the framework of possibility theory described in the first two sections of this paper. Its weighted version is an extension of possibilistic logic, but it has arguably a limited expressive power. On the other hand, the multi-agent logic  $\text{GBDL}^n$  is a graded extension of the full-fledged multimodal logic  $\text{KD}^n$  that is very expressive, but is arguably overexpressive as it contains formulas that can be made hardly intuitive sense, and its semantics is much richer than the framework of possibility theory. So, there is a need for more research on the bridge between modal and possibilistic logics. The logic  $\text{GBDL}^n$ :

- allows for objective formulas while MBDL does not. What is their role and can we dispense with them?
- allows for introspective formulas of the form  $N_1^{\geq \alpha_1} N_1^{\geq \alpha_2} \varphi$  that are not part of the setting of possibility theory. How to make sense of them?
- presupposes that the epistemic state of an agent depends on the (objective) state this agent is in, which is not part of the formal framework described in the first sections. It enables standard techniques in modal logic to be applied, but it is not always easy to interpret.



More work is needed to come up with an epistemic logic framework which is at the same time expressive enough for our purpose, and where both semantic and syntactic aspects remain under control.

## References

1. Banerjee, M., Dubois, D.: A simple modal logic for reasoning about revealed beliefs. In: Sossai, C., Chemello, G. (eds.) ECSQARU 2009. LNCS, vol. 5590, pp. 805–816. Springer, Heidelberg (2009)
2. Benferhat, S., Dubois, D., Kaci, S., Prade, H.: Bipolar possibility theory in preference modeling: Representation, fusion and optimal solutions. *Information Fusion* 7, 135–150 (2006)
3. Blackburn, P., de Rijke, M., Venema, Y.: *Modal Logic*. Cambridge University Press (2001)
4. Boutilier, C.: Towards a logic for qualitative decision theory. In: *Principles of Knowledge Representation and Reasoning: Proc. of the 5th Int. Conf. (KR 1994)*, pp. 75–86. AAAI Press (1994)
5. Casali, A., Godo, L., Sierra, C.: A graded BDI agent model to represent and reason about preferences. *Artif. Intell.* 175, 1468–1478 (2011)
6. Dastani, M., Lorini, E.: A logic of emotions: from appraisal to coping. In: *Proc. of AAMAS 2012*, pp. 1133–1140. ACM Press (2012)
7. Dubois, D., Hajek, P., Prade, H.: Knowledge-driven versus data-driven logics. *J. of Logic, Language, and Information* 9, 65–89 (2000)
8. Dubois, D., Prade, H.: Resolution principles in possibilistic logic. *Int. J. Appr. Reas.* 4, 1–21 (1990)
9. Dubois, D., Prade, H.: Epistemic entrenchment and possibilistic logic. *Artificial Intelligence* 50, 223–239 (1991)
10. Dubois, D., Prade, H.: Possibility theory: qualitative and quantitative aspects. In: Gabbay, D., Smets, P. (eds.) *Quantified Representation of Uncertainty and Imprecision. Handbook of Defeasible Reasoning and Uncertainty Management Systems*, vol. 1, pp. 169–226. Kluwer (1998)
11. Dubois, D., Prade, H.: From Blanchés hexagonal organization of concepts to formal concept analysis and possibility theory. *Logica Universalis* 6(1), 149–169 (2012)
12. Dubois, D., Prade, H.: Possibilistic logic: a retrospective and prospective view. *Fuzzy Sets and Systems* 144, 3–23 (2004)
13. Fariñas del Cerro, L.: Resolution modal logic. *Logique et Analyse* 110–111, 153–172 (1985)
14. Fagin, R., Halpern, J.Y., Moses, Y., Vardi, M.Y.: *Reasoning about Knowledge*. MIT Press (1995)
15. Gratch, J., Marsella, S.: A domain independent framework for modeling emotion. *Cognitive Systems Research* 5(4), 269–306 (2004)
16. Goldszmidt, M., Pearl, J.: Qualitative probability for default reasoning, belief revision and causal modeling. *Artificial Intelligence* 84, 52–112 (1996)
17. Hintikka, J.: *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell University Press (1962)
18. Hume, D.: *A Treatise of Human Nature*. In: Selby-Bigge, L.A., Nidditch, P.H. (eds.). Clarendon Press, Oxford (1978)
19. Lang, J., van der Torre, L., Weydert, E.: Hidden Uncertainty in the Logical Representation of Desires. In: *Proc. of Int. Joint Conf. on Artificial Intelligence*, pp. 685–690. Morgan Kaufmann (2003)
20. Lang, J., van der Torre, L., Weydert, E.: Utilitarian Desires. *J. of Autonomous Agents and Multi-Agent Systems* 5, 329–363 (2002)

21. Lorini, E.: A dynamic logic of knowledge, graded beliefs and graded goals and its application to emotion modelling. In: van Ditmarsch, H., Lang, J., Ju, S. (eds.) LORI 2011. LNCS, vol. 6953, pp. 165–178. Springer, Heidelberg (2011)
22. Lorini, E., Prade, H.: Strong possibility and weak necessity as a basis for a logic of desires. In: Godo, L., Prade, H. (eds.) Working Papers of the ECAI 2012 Workshop on Weighted Logics for Artificial Intelligence (WL4AI 2012), Montpellier, pp. 99–103 (August 28, 2012)
23. Meyer, J.J., van der Hoek, W.: Epistemic Logic for AI and Computer Science. Cambridge University Press (1995)
24. Ortony, A., Clore, G., Collins, A.: The Cognitive Structure of Emotions. Cambr. Univ. Pr. (1988)
25. Sahlqvist, H.: Completeness and correspondence in the first and second order semantics for modal logic. In: Proc. of 3rd Scandinavian Logic Symp., pp. 110–143. North-Holland (1975)
26. Shackle, G.L.S.: Decision, Order, and Time in Human Affairs. Cambridge Univ. Press (1961)
27. Spohn, W.: Ordinal conditional functions: a dynamic theory of epistemic states. In: Causation in Decision, Belief Change and Statistics, pp. 105–134. Kluwer (1988)
28. Zadeh, L.A.: Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets & Syst. 1, 3–28 (1978)